Earth Syst. Dynam. Discuss.,
https://doi.org/10.5194/esd-2018-13-RC2, 2018

Earth System Dynamics Discussions

Interactive comment on "Using Network Theory and Machine Learning to predict El Niño" *by* Peter D. Nooteboom et al.

Anonymous Referee #2

Received and published: 13 April 2018

Overall I think this is valuable and important work, but I think there could be more clarity in the writing. It tends to read as a long sequence of sentences rather than a narrative that walks the reader through the steps of the analysis. At the end, I'm left slightly confused as to (i) how did you use the CZ model; did you actually learn something from that that helped analyze the real world, (ii) how you decided on the specific input variables (rather than what sounds like a jumbled mess of exploring a wide variety of different concepts that might have some relevance), (iii) to what extent your improvement in prediction is actually related to ML/ANN versus having identified good predictive variables (e.g., could you have identified a linear model that used those variables and obtained a good prediction? Were the ultimate relationships "learned" by the ANN between inputs and output actually notably nonlinear?), and (iv) it would help

to have a single final plot showing rms error vs prediction horizon as compared with the current methods.

1. P2, 1st line, not quite sure how to define "intuition and creative thinking", nor (more importantly) why this is relevant here.

2. P2, par lines 3-11, this seems a bit awkwardly worded. It isn't a binary choice between many layers and inputs and "simpler", but rather a continuum of choices with an inherent trade-off. Using more layers and input variables means you can rely more on the algorithm to figure out what matters at the expense of needing to train it on more training data, and the fewer variables/layers one uses the less training data might be required but the more that forces the user to make intelligent choices for input variables rather than relying on the algorithm to do so.

3. The choices in Section 2.3 are not well motivated (that is, why are these the relevant choices to feed into the ANN, and what else did you try?) This section could benefit from a couple of introductory sentences that describe the goal of the section, and the broad overview of the ideas of the section.

4. Why is it adequate to have all of the memory embedded in the linear part of the model?

5. For that matter, not entirely obvious to me, since you are using ML to predict the nonlinear terms anyway, whether the ML can also predict the linear (but dynamic) part without any extra effort, or for that matter the nonlinear and dynamic part. Did you try different things and conclude you didn't have enough training data to converge, and kept simplifying, or did you just guess what might work and then it did? I didn't go back and read Hibon and Evgeniou, but it would seem like the question of how to simplify what the ML is actually learning is case dependent rather than absolute. Some more motivation here is required (and at a minimum you should clarify what is meant by "more stable" and provide a few more words of intuition as to why this reduces the risk of a bad prediction.)

6. Extra plus sign in eqn 13 and 14. Also, shouldn't the summation on the second term start at d+1 (otherwise, the j=1 in the second term and the i=1 in the first term are identical, and you have a standard ARMA model rather than an ARIMA model). (Also, don't recall if you said why you were using ARIMA rather than ARMA?)

7. P7, L19-20, why would including past El Nino and La Nina information reduce prediction skill?

8. P8, L1, I'd have just thought the choice of lead time is like a choice of different variables, that there's nothing wrong with including the same variable at different times as part of the input.

9. P8, L17, "generally" as in, "in this paper", or "generally" as in "in most research"?

10. Section 3.1, any reason why you only used 45 years of ZC output? Why not use a few thousand years of output? (I ran it for that long quite a long time ago, so I know it isn't a computational challenge to do.)

11. Also, section 3.1, you might want to say up front a bit more about motivation – are you trying to learn from ZC which variables are best to use, or ultimately comparing predictive capability on ZC vs the real world, or get a good initial estimate of ANN weights from ZC so that you don't have to converge as much when you apply to the real world? These are all possible goals, but other than the second one, may be problematic if the physics in ZC doesn't match the real world physics (and while with their original parameter choices the equilibrium point in ZC is unstable with a chaotic self-sustained response, I think the general consensus now is that the real world isn't exhibiting chaos but rather stochastically forced response of a damped stable system). This is similar to the comment on Section 2.3; it would be helpful to have a few additional sentences that talk about where you're going with a section, why is it here, what are you hoping to learn, and what the structure of the section is. (I note subsequently that you never actually look at the predictability of CZ model, improvement thereof with ANN, and you also don't use the same variables in the real world analysis. . . can you be clear as to

why this section is here and what you learned? Is it here just because you spent a lot of time on it and figure that should be documented somewhere, or is it essential to motivate the analysis of the real world?)

12. P10, L2, I think what you mean here is something like "when the ENSO index changes from increasing to decreasing (peak El Nino) or from decreasing to increasing (peak La Nina)"? (The wording is a bit unclear to me.) Similarly line 7, refer to the derivative of the ENSO index, rather than the derivative of ENSO. . . (to me, "ENSO" refers to the overall dynamic phenomenon, which isn't a thing that has a sign or a derivative).

13. Section 4, rather than just focusing on a few things like 2010 (which is cherry-picked as a year where the default scheme does badly), and a few prediction horizons, one thing that would help evaluate this method would be a single plot of rms prediction error versus time for the two methods (that is, for any month once you have sufficient past data, do the N-month prediction for every N up to a year or more using both methods, and then over this big set of month N predictions, what's the rms error?) This would also be a great way to compare your ARIMA alone with ARIMA + ANN.

14. P14, L11, what do you mean by "best-performing"? What metric? Does that mean that adding more neurons made it worse? Or do you just mean that adding more neurons didn't make it better?

15. P15, L4, why compare the two methods at different lags instead of the same lag?

16. P15, L7, doesn't this contradict the abstract?

17. P15, L14, I'm confused by this sentence – you do a better job at predicting things 1 year in advance than 6 months?

18. Also, I must have missed something; I thought you'd already picked the set of input variables, and now it sounds like you are only using a subset, and a different subset for each prediction horizon. Overall, this sounds incredibly fragile. You do a lot of work to

pick a few really good input variables, and any time you change the time horizon you might need to change those, and change the number of neurons. . . I thought the whole point of ANN was the ability to be lazy and let the algorithm do all the work for you!

19. P15, L15-16, again, I'm a bit confused. . . why do we need to maintain a whole ensemble of different ANN structures? This doesn't converge to something with enough neurons? Also, Figure 11, am I interpreting this right that you found a bunch of possible ANN structures that outperform the ones in Figure 9? (Sorry, I'm totally lost at this point so this might be off-base and simply imply some insufficient description.) Why not go back and redo Fig 9 with the better ANN structure? This entire section reads a bit as a collection of odds and ends of results rather than as a post-facto summary.

Interactive comment on Earth Syst. Dynam. Discuss., https://doi.org/10.5194/esd-2018-13, 2018.